



Comparison of genetic algorithm systems with neural network and statistical techniques for analysis of cloud structures in midlatitude storm systems

J.A. Parikh^{a,*}, J.S. DaPonte,^b J.N. Vitale,^a G. Tselioudis^{b,1}

^a *Computer Science Department, Southern Connecticut State University, New Haven, Connecticut 06515, USA*

^b *Columbia University and NASA Goddard Institute for Space Studies, New York, NY 10025, USA*

Abstract

Cloud analyses provide information which is vital to the detection, understanding and prediction of meteorological trends and environmental changes. This paper compares statistical, neural network and genetic algorithm methods for recognition and tracking of midlatitude storm clouds in sequences of low-resolution cloud-top pressure data sets. Regions of interest are identified and tracked from one image frame to the next consecutive frame in an eight-frame sequence. Classification techniques are used to determine the relationships between regions of interest in consecutive time frames. A genetic algorithm procedure is then used to revise classifier outputs to ensure that consistency constraints are not violated. © 1997 Elsevier Science B.V.

Keywords: Cloud tracking; Neural network; Genetic algorithm

1. Introduction

The long-term goal of our research is to develop automatic techniques to characterize changes in the multiscale properties of storm systems in the northern midlatitudes. In the midlatitude regions, clouds are primarily formed by synoptic scale storm systems that are often referred to as midlatitude cyclones. Cyclones play a crucial role in determining the Earth's radiative balance and are a major component in the global hydrological cycle. Accurate characterization of the distribution and properties of cloud

systems is essential for predictions of global climate change. By diminishing uncertainties in cloud formation mechanisms, uncertainties in General Circulation Model (GCM) predictions of global change can be greatly reduced.

The focus of our current research is on the recognition and tracking of midlatitude storm systems. Previous approaches to tracking cloud motion have included cross correlation, cost function, feature identification, optical flow, sea-level pressure and geopotential height field methods (Vega-Riveros and Jabbour, 1989; Schmetz et al., 1987; Endlich and Wolf, 1981; Bolla et al., 1995; Cohen and Herlin, 1996; Murray and Simmonds, 1991; Konig et al., 1992). In our approach sequences of low-resolution cloud-top pressure data sets were processed to obtain regions of interest using thresholding followed by

* Corresponding author. E-mail: parikh@scsu.ctstateu.edu.

¹ E-mail: ccgxt@giss.nasa.gov.

connected component labeling. Features were extracted for pairs of regions of interest in consecutive time frames. Paired regions of interest were then classified into one of four classes – a correspondence class, a merge class, a split class or a reject class. A genetic algorithm procedure was then applied to neural network classifier outputs to produce a revised and consistent set of paired regions for consecutive time frames.

2. Regions of interest

The data sets used in this study were produced by the International Satellite Cloud Climatology Project (ISCCP) (Rossow and Schiffer, 1991) and extracted from 3-hourly cloud products available over the Web from NASA Langley Research Center EOSDIS Archive Center. Data sets for two days in the year 1992, on which midlatitude cyclones had been observed and reported in the literature (Brown and Zeng, 1994), were selected for this study. From each of the selected data sets, cloud-top pressure data derived from “window” infrared ($IR = 11 \mu m$) imagery were extracted for latitudes from $15^\circ N$ to $75^\circ N$. Each value represents information for an area of approximately $2.5^\circ \times 2.5^\circ$. Values for cloud-top pressures range from 50 to 1000 mb with values in the range from 50 to 440 mb denoting high-level clouds and values in the range from 440 to 680 mb denoting middle-level clouds.

The first step in the construction of the regions of interest was to threshold the data sets at the ISCCP cloud-top pressure threshold of 440 mb for high-level clouds. The initial regions were then allowed to expand until either a significant change in pressure (specified as 40 mb) was encountered or the regions had extended to the 560 mb range (the mid-point for middle-level clouds). Each region was then labeled using a standard connected component algorithm for 8-neighbor connectedness.

3. Feature selection

After identification of regions of interest in each of the eight time frames for the training and test data sets, a total of twenty-nine features were extracted for each pair of regions in consecutive time frames.

The features represented differences in area, position, cloud-top pressure and texture. Texture features calculated for each region were the sum and difference histogram texture features described in Bankert (1994).

Stepwise discriminant analysis was applied to the feature set to determine which features could distinguish among the categories of splitting, merging, direct correspondence and reject. The analysis yielded five key features. The sum of the absolute values of the differences between minimum latitude and maximum latitude – a measure of the relative stability of cloud systems with respect to north-to-south movement – was found to be the best feature. The other four features selected by stepwise discriminant analysis were the features representing difference in areas between regions in consecutive time frames, the absolute value of the difference in areas, the absolute value of the difference in the centroid longitudes, and the absolute value of the difference in maximum cloud-top pressure. This reduced feature set was then used for the classification experiments described in the next section.

4. Classification

Linear discriminant analysis, nearest neighbor ($k = 3$) analysis, statistically-based neural network models from the SAS Institute, standard backpropagation neural network classifiers, and genetic algorithm neural network classifiers were designed and developed for classification of region pairs. The classifiers were designed on a subset of the training set region pairs and tested on the complete set of 744 test set region pairs. The only difference between the training set and the complete set of all possible training set region pairs was the exclusion of some pairs designated as “Reject”. All region pairs ($region_1, region_2$) in both training and test sets were classified manually into either a Reject (R) class, a Correspondence (C) class, a Split (S) class or a Merge (M) class.

The architectures of the SAS neural network, the backpropagation neural network and the genetic algorithm neural network were identical. All the neural network classifiers were trained and tested on the five normalized features selected by stepwise discriminant analysis. Each network had five input

Table 1
Test set classification results

Classifier	Accuracy	Weighted Kappa	Confidence interval
Linear discriminant	78.9%	0.534	0.464–0.605
Nearest neighbor	75.9%	0.415	0.346–0.485
SAS neural net	96.6%	0.899	0.858–0.940
BP neural net	96.2%	0.876	0.829–0.924
GA neural net	94.5%	0.837	0.785–0.890
Constraint-handling GA	97.0%	0.907	0.866–0.949

nodes, two hidden nodes and two output nodes. The desired values at the output nodes for each of the four classes were (0.1,0.1) for class R, (0.9,0.9) for class C, (0.9,0.1) for class S, and (0.1,0.9) for class M. The values of the learning rate and momentum for the standard backpropagation network were set to 0.02 and 0.8, respectively. The genetic algorithm network used a string length of 18 to correspond to the 18 real-valued weights of the neural network and terminated with a stopping criterion of 10,000 iterations. The genetic algorithm library PGAPack (Levine, 1996) from Argonne National Laboratory was used for development of all genetic algorithms used in this study.

The results of each of the classification techniques are summarized in Table 1. The overall classification accuracy, the value of the Weighted Kappa Statistic and the 95% confidence interval for the Kappa Statistic are given for each of the classifiers. The Kappa Statistic (Cohen, 1968) measures the agreement between a particular classification scheme and the “ground truth” for the patterns. In general, the interpretation of Kappa can be considered as poor agreement for values between 0.0 to 0.4, fair to good agreement for values between 0.4 and 0.75, and excellent agreement for values between 0.75 to 1.0.

The ability of the neural network classifiers to distinguish direct correspondence of regions and reject regions was primarily responsible for their high

overall classification accuracy. It can be seen, however, from the confusion matrix results shown in Tables 2 and 3 that the Nearest Neighbor classifier performed better than the SAS Neural Net for classification of merges (68% versus 36%). Linear discriminant analysis correctly identified all of the direct correspondences and 64% of merges but only 45% of splits and 79% of rejects. These results suggest that future investigations should include statistical pattern recognition methods as well as neural network approaches.

5. Constraint-handling genetic algorithm

The genetic algorithm constraint-handling procedure was designed to improve the classification output of the neural network classifiers by ensuring that logical constraints on classification outputs were satisfied. Each individual or chromosome in the population was an integer-valued string representing possible region pair classifications. The string length of each individual was equal to the total number of region pairs. The value of each allele within a chromosome ranged from 0 to 3 denoting a classification category of R, C, S or M, respectively. If the value for a given allele was zero, which represented a classification of R for reject, the sum-squared error between the neural network outputs for the corre-

Table 2
Nearest neighbor ($K = 3$) analysis test set results

Class	No. Patterns	C	M	S	R
C	39	21 (54%)	13	3	2
M	22	4	15 (68%)	1	2
S	22	2	0	16 (73%)	4
R	661	14	70	82	495 (75%)

Table 3
SAS Neural Network Test Set Results

Class	No. Patterns	C	M	S	R
C	39	39 (100%)	0	0	0
M	22	8	8 (36%)	0	6
S	22	2	0	17 (77%)	3
R	661	0	1	5	655 (99%)

sponding region pair and the neural network output values for class R was calculated. The fitness measure for any chromosome which did not violate any constraints was just the total sum-squared error.

After a specified number of iterations, a penalty in the fitness function was triggered by violation of any of the following constraints:

1. No region can correspond directly to more than one region in either a previous or a consecutive time frame.
2. A region that splits must split into two or more regions in the next time frame.
3. If there is a merge from time frame 1 to time frame 2, then there must be two or more regions in time frame 1 which merge into a given region in time frame 2.
4. A given region cannot both directly correspond to a region in a consecutive time frame and split into another region in the same time frame.
5. A given region cannot both directly correspond to a region in a consecutive time frame and merge into another region in the same time frame.

Experiments using the SAS neural net outputs as inputs were run for each pair of consecutive time frames in the test set. Parameter values were 5,000 for maximum number of iterations, 0.3 for crossover probability, 500 for population size, 1,000 for constraint violation penalty, and 1,000 for number of iterations before application of any penalty. There were eight instances of constraint violation in the test set classification results produced by the SAS neural network. As a result of applying the genetic algorithm constraint-handling procedure to the SAS neural network outputs, no constraints were violated and overall classification accuracy was improved from 96.6% to 97.0% (see Table 1).

The ability of the genetic algorithm procedure to converge to nearly optimal solutions which satisfied the given constraints was highly impressive. The convergence of the procedure did depend on waiting until a certain number of iterations had passed before applying penalties. An excellent discussion of constraint-handling techniques in evolutionary computation methods and, in particular, of penalty function methods can be found in Michalewicz (1995). For our application, we investigated parameter dependence by, for example, increasing crossover probability from 0.3 to 0.5 or by reducing population size

from 500 to 200. These changes did not significantly affect results.

6. Conclusions

The experimental results of the classification and genetic algorithm procedures on the cloud-top pressure data sets demonstrate the potential of neural network and genetic algorithm procedures to identify and track cloud systems of interest. As expected, the major classification problems arose in detection of the cases when either two or more cloud systems merged together or when a cloud system split into two or more regions. Development of an operational method to determine distribution of and changes in high-level, optically thick clouds in the northern midlatitudes requires additional meteorological and satellite data set collection, further refinement of processing techniques, and more extensive testing on a variety of meteorological data sets. Future research plans include integration of optical thickness data sets with cloud-top pressure data sets, collection and analysis of National Meteorological Center data sets for building “ground-truth” training sets for classification procedures, verification of the reproducibility of classification results, and continued development of constraint-handling evolutionary procedures for tracking cloud systems.

Acknowledgements

The authors would like to express their appreciation to William B. Rossow for his guidance and contributions to this research. The ISCCP D1 data sets which were used in this study originated at NASA Goddard Institute for Space Studies and were downloaded from the webserver at the NASA Langley Research Center EOSDIS Distributed Active Archive Center. The authors would like to acknowledge the SAS Institute, Inc. of Cary, NC for the provision of macros implementing various neural network models using the SAS product.

Discussion

Serpico: If I understand correctly, you selected features before experimenting with all these classifiers.

Can you add some detail about how you did feature selection?

Parikh: We actually did several experiments with different combinations of features. What I showed here are the results for the features which were selected by using stepwise discriminant analysis. We did use feature combinations not only of five features, but also of six and seven features. We used neural networks of different sizes. We tried different seeds. We found, in general, that the nearest neighbour technique and the linear discriminant analysis picked up the splits and the merges. The neural network classifiers on the whole were doing better for overall accuracy.

Hancock: I was interested in your genetic algorithm strategy in which you removed all but one of the solutions in the population, before moving on to the next generation. Does that give the algorithm any chance to perform crossover if you remove all but the fittest solution from the population?

Parikh: I first experimented with a very high mutation rate, eliminating crossover almost completely. That did not give very good results. In the final algorithm, crossover was performed with a probability of 0.3.

Hancock: Could you provide some rationale as to why that particular choice of parameters worked so well?

Parikh: I mentioned that I want to try some other techniques and look further into exactly what is happening with these constraint satisfaction prob-

lems. I did have problems with constraint satisfaction finding global minima without proper choice of parameters.

References

- Bankert, R.L., 1994. Cloud classification of AVHRR imagery in maritime regions using a probabilistic neural network. *J. Appl. Meteor.* 33, 909–918.
- Bolla, R., Marchese, M., Nobile, C., Zappatore, S., 1995. Prediction of short-term evolution of cloud formations based on meteosat image sequences. In: *Proc. ICIAP'95*, pp. 677–682.
- Brown, R.A., Zeng, L., 1994. Estimating central pressures of oceanic mid-latitude cyclones. *Appl. Meteor.* 33 (9).
- Cohen, I., Herlin, I., 1996. Optical flow and phase portrait methods for environmental satellite image sequences. In: *Proc. ECCV'96*.
- Cohen, J., 1968. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol. Bull.* 70, 213–220.
- Endlich, R.M., Wolf, D.E., 1981. Automatic cloud tracking applied to goes and meteosat observations. *J. Appl. Meteor.* 20 (3), 309–319.
- König, W., Sausen, R., Sielmann, F., 1992. Objective identification of cyclones in GCM simulations. *J. Climate* 6, 2217–2231.
- Levine, D., 1996. Users guide to the PGAPack parallel genetic algorithm library.
- Michalewicz, Z., 1995. A survey of constraint handling techniques in evolutionary computation methods. In: *Evolutionary Programming IV*, pp. 135–155.
- Murray, R.J., Simmonds, I., 1991. A numerical scheme for tracking cyclone centres from digital data. *Aust. Met. Mag.* 39, 155–180.
- Rosow, W.B., Schiffer, R.A., 1991. ISCCP cloud data products. *Bull. Am. Meteor. Soc.* 72 (1), 2–20.
- Schmetz, J., Holmlund, K., Hoffman, J., Strauss, B., Mason, B., Gaertner, V., Koch, A., Berg, L.V.D., 1987. Operational cloud-motion winds from meteosat infrared images. *J. Appl. Meteor.* 32, 1206–1225.
- Vega-Riveros, J.F., Jabbour, K., 1989. Review of motion analysis techniques. *IEE Proc.* 136 (6).